

PAT-NO: **JP401191200A**

DOCUMENT-IDENTIFIER: JP 01191200 A

TITLE: VOICE RECOGNIZING DEVICE

PUBN-DATE: August 1, 1989

INVENTOR-INFORMATION:

NAME

AKAHA, MASATERU

ASSIGNEE-INFORMATION:

NAME	COUNTRY
SONY CORP	N/A

APPL-NO: JP63015560

APPL-DATE: January 26, 1988

INT-CL (IPC): G10L003/00, G06F003/16 , G06F015/62

US-CL-CURRENT: 704/247, 704/273

ABSTRACT:

PURPOSE: To quickly execute a response of a voice recognition by discriminating a fingerprint of a talker, discriminating a fact that its talker is a registered talker who has registered a voice pattern and a fingerprint and executing pattern matching of the voice pattern which this registered talker has registered and an input voice pattern.

CONSTITUTION: The title device is provided with a standard voice pattern memory 9, a pattern matching means 6, a fingerprint reading means 7, a fingerprint registering means 12 and a fingerprint discriminating means 8. Plural pieces of common specific words by plural talkers, and original words of each talker are registered in the standard voice pattern memory 9. Also, a fingerprint of a talker who has executed a word registration is registered in the fingerprint registering means 12. Subsequently, whether the fingerprint

which has been read by the fingerprint reading means is the fingerprint which has been registered in the fingerprint registering means 12 or not is discriminated by the fingerprint discriminating means 8, and when it is the registered fingerprint, only a voice pattern of the talker having this fingerprint is supplied to the pattern matching means 6. In such a way, the responsiveness of a voice recognition of the registered talker can be improved.

⑫ 公開特許公報 (A) 平1-191200

⑬ Int. Cl.⁴

G 10 L 3/00
G 06 F 3/16
15/62
G 10 L 3/00

識別記号

3 0 1
3 2 0
4 6 5
3 0 1

府内整理番号

D-8842-5D
G-7341-5B
A-6615-5B
C-8842-5D

⑭ 公開 平成1年(1989)8月1日

⑮ 発明の名称 音声認識装置

⑯ 特願 昭63-15560

⑰ 出願 昭63(1988)1月26日

⑱ 発明者 赤羽正照 東京都品川区北品川6丁目7番35号 ソニー株式会社内

⑲ 出願人 ソニー株式会社 東京都品川区北品川6丁目7番35号

⑳ 代理人 弁理士 松隈秀盛

明細書

発明の名称 音声認識装置

特許請求の範囲

複数の話者による共通の複数個の特定単語と、各話者独自の単語とが登録される根拠音声パターンメモリと、

根拠音声パターンと入力音声パターンとのマッチングを行なうパターンマッチング手段と、

指紋読取手段と、

上記根拠音声パターンメモリに単語登録を行なった話者の指紋が登録される手段と、

上記指紋読取手段によって読み取られた指紋が上記指紋登録手段に登録された指紋であることを判別したら、この読み取られた指紋を有する話者により登録された根拠音声パターンのみを上記根拠音声パターンメモリから、上記パターンマッチング手段に供給する手段とを備えた音声認識装置。

発明の詳細な説明

(産業上の利用分野)

この発明は、音声認識装置に関する。

〔発明の概要〕

この発明は、話者の個人特有のものである指紋の判別を行なって、その話者が、音声登録ならびに指紋登録がなされた登録話者であるか否かを判別し、その話者が登録話者の場合には、その登録話者が登録した音声パターンと入力音声パターンとのパターンマッチングを行ない音声認識を行なうようにし、話者が登録話者であるか否かの判別を容易に、そして確実にできるようにしたもので、その登録話者の音声認識の応答性を良くするとともにその登録話者が登録した秘密情報の専用使用の保持に好適なものである。

〔従来の技術〕

音声は時間軸に沿って変化する現象で、スペクトル・パターンが刻々と変化するよう音声を発声することによって、固有の単語や言葉が生まれる。この人間が発声する単語や言葉を自動認識する技術が音声認識であるが、人間の聴覚機能に匹敵するような音声認識を実現することは現在の

ところ至難のことである。このため、現在使用されている音声認識の殆んどは、一定の使用条件下で、認識対象単語の標準パターンと入力パターンとのパターンマッチングを行なうことによりなす方法である。

第5図はこの音声認識装置の概要を説明するための図で、マイクロホン(1)よりの音声入力が音響分析回路(2)に供給される。この音響分析回路(2)では入力音声パターンの特徴を表わす音響パラメータが抽出される。そして、この音響分析回路(2)よりの音響パラメータは、例えばスイッチからなる切換回路(3)に供給される。この回路(3)のスイッチが端子A側に切り換えられるときは登録時で、音響パラメータが認識パラメータとして、標準音声パターンメモリ(4)にストアされる。つまり、音声認識に先だって話者の音声パターンが標準パターンとして、このメモリ(4)に記憶される。

一方、切換回路(3)のスイッチが端子B側に切り換えられるときは認識時である。そして、この認識時は、音響分析回路(2)からのそのときの入力音

声の音響パラメータが入力音声パターンメモリ(5)に供給されて一時ストアされる。そして、この入力音声の音響パラメータと、標準音声パターンメモリ(4)から読み出された複数の認識対象単語の標準音声パターンの音響パラメータとが、パターンマッチング回路(6)で比較される。そして、このパターンマッチング回路(6)にて、入力音声の音声パターンに最もよく似た標準音声パターンが検出される。

さて、同じ言葉であっても、それを発音した場合、話し方の違いや発声器官の違いのため、音声の音響的性質は発声者によって、大きく異なる。したがって、上述した音声認識装置の標準音声パターンメモリ(4)に、ある話者の音声パターンを標準パターンとして登録した場合、その話者個人専用となり、複数の話者、つまり不特定話者に対しては、単語の判別が困難となる。

そこで、一つの単語に対して、複数の話者の音声パターンを標準パターンとしてメモリに記憶させ、そして、この標準パターンと入力音声バター

ンとを比較することによって、話者方向の判別度下げ、登録単語方向の判別度を上げて、単語登録していない話者にも、対応できるようにして、複数の話者を対象とした音声認識が行なわれている。

(発明が解決しようとする課題)

ところで、上述したように、一つの単語に対して、複数の話者の音声パターンを標準パターンとして記憶させ、話者方向の判別度を下げれば、複数の話者つまり不特定話者に対する登録単語の判別度は高くなるが、話者の判別はうまくできなくなってしまう。この場合、例えば、ある特定の者が機密情報を音声認識装置に記憶したとすると、その者の音声であるかどうかが判別できないので、その機密情報が、他の者にも使用されてしまう危険性がある。

この危険性を避けるためには、例えば暗号を設定し、この暗号を音声にて、音声認識装置に入力することが考えられる。しかし、この暗号の発声時には、暗号が他人に聞きとられる可能性が高く、

秘密性の点で適切ではない。また、例えば暗証番号等をキーにて、音声認識装置に入力することも考えられるが、これも、秘密性の点で、確実なものではなく、また、キー入力の煩しさも伴ない、不適当である。

(課題を解決するための手段)

この発明は、標準音声パターンメモリ(4)と、パターンマッチング手段(6)と、指紋読取手段(7)と、指紋登録手段(8)と、指紋判別手段(9)とを備える。

(作用)

複数の話者による共通の複数個の特定単語と、各話者独自の単語とが標準音声パターンメモリ(4)に登録される。そして、単語登録を行なった話者の指紋が指紋登録手段(8)に登録される。指紋読取手段(7)で読み取られた指紋が、上記指紋登録手段(8)に登録された指紋であるか否かが、指紋判別手段(9)によって判別され、登録された指紋であれば、この指紋を有する話者の音声パターンのみが、バ

ターンマッチング手段⑥に供給される。

(実施例)

第1図は、この発明の一実施例のブロック図であり、例えば電話機の例である。なお、図において、第5図例と同等なものには同一符号を付してある。

図において、Ⓐはハンドセットの送話器のマイクロホン、Ⓑは指紋読取手段、Ⓒは指紋判別手段、Ⓓは指紋登録メモリ、Ⓔは標準音声パターンメモリ、Ⓕはそのアドレス制御回路、Ⓖはコントローラである。

標準音声パターンメモリ⑤は、例えば第2図に示すように、話者Ⅰ、Ⅱ、Ⅲ毎の単語A、B、Cの音声パターンが登録された共通パターンメモリ⑨A)と、話者、Ⅰ、Ⅱ、Ⅲ毎の秘密情報である独自の単語X、Y、Zの音声パターンが登録された専用パターンメモリ⑨B)とからなる。なお、共通パターンメモリ⑨A)には、秘密ではないが、各話者独自の単語が登録してあってもよい。この標準

音声パターンメモリ⑨に音声を登録する場合の動作を以下に説明する。

第3図は、音声登録の動作のフローチャートを示す図である。

同図において、ステップ(201)にて、共通パターンメモリ⑨A)への登録か、それとも専用パターンメモリ⑨B)への登録かが判断される。共通パターンメモリ⑨A)への登録であれば、ステップ(202)にて、例えば、ハンドセットに設けられた指紋読取手段Ⓑによって指紋認識が行なわれ、第1図に示す指紋登録メモリⒹにその指紋が登録される。そしてステップ(205)にて、入力音声パターンが共通パターンメモリ⑨A)に記憶される。また、専用パターンメモリ⑨B)への登録であれば、ステップ(203)にて、指紋の認識が行なわれ、第1図に示す指紋登録メモリⒹにその指紋が登録される。そして、ステップ(204)にて、入力音声パターンが専用パターンメモリ⑨B)に記憶される。

次に、音声を認識する場合の動作を以下に説明する。

第1図において、まず、話者が電話機のハンドセット(図示せず)を取り上げる。すると、このハンドセットには、指紋読取手段Ⓑが備えられており、話者がハンドセットを握ることによって、この話者の指紋が指紋読取手段Ⓑに読み取られ、そして、一時的に指紋を示す信号がこの指紋読取手段Ⓑに記憶される。そして、話者が共通モードを示す音声、例えば「キョウツウ」をマイクロホンⒷに向かって発声すると、音響分析回路⑧そして、切換回路⑩を介して、発声された音声に対応した信号が入力音声パターンメモリ⑥に供給され、一時この入力音声パターンメモリ⑥に記憶される。そして、共通パターンメモリ⑨A)から読み出された音声パターンと、入力音声パターンとがパターンマッチング回路⑦に供給され、入力された単語、すなわち「キョウツウ」が認識される。そして、この認識された単語に対応した信号がコントローラⒼに供給される。すると、コントローラⒼは指紋読取手段Ⓑおよび指紋判別手段Ⓒに共通モードであることを示す信号を供給する。すると、指紋

読取手段Ⓑに一時的に記憶された、話者の指紋を示す信号がこの指紋読取手段Ⓑから指紋判別手段Ⓒに供給される。すると、この指紋判別手段Ⓒは供給された指紋を示す信号が指紋登録メモリⒹに指紋が登録され、そして標準音声パターンメモリ⑤に音声パターンが登録された登録話者の指紋を示すものかどうかを判別する。そして、指紋判別手段Ⓒは、供給された指紋を示す信号が上記登録話者を示すものであれば、共通モードであることを示し、そしてその登録話者を示す信号をアドレス制御回路Ⓖに供給する。すると、このアドレス制御回路Ⓖは、その登録話者の音声パターンが登録されたアドレスを示す信号を共通パターンメモリ⑨A)に供給する。すると、この共通パターンメモリ⑨A)に登録された、その登録話者のみの音声パターンが選択される。つまり、例えば、第2図に示す、共通パターンメモリ⑨A)の例えば、登録話者Ⅰの単語A、B、Cの音声パターン、ⅠA、ⅠB、ⅠCが選択され、これらが、パターンマッチング回路⑦に供給されるように制御される。ま

た、もし、指紋判別手段④に供給された、指紋を示す信号が、指紋登録された登録話者のものでなければ、例えば話者Ⅰの音声パターンだけでなく、共通パターンメモリ(9A)に登録された話者全ての音声パターンが、例えば、第2図の共通パターンメモリ(9A)の話者Ⅰ、Ⅱ、Ⅲの音声パターンがパターンマッチング回路⑥に供給されるように制御される。

そして、次にマイクロホン⑪に、例えば「ツウワ」という音声が供給されると、この「ツウワ」を示す音声パターンと、共通パターンメモリ(9A)からの音声パターンとが、パターンマッチング回路⑥にて比較される。そして、このパターンマッチング回路⑥で、例えば上述した「ツウワ」という単語が認識される。すると、この「ツウワ」という単語に対応した信号が、パターンマッチング回路⑥から、コントローラ⑩に供給される。すると、このコントローラ⑩は、電話機が通話可能な状態となるように、電話機を制御する。そして、次に、話者が音声等によって例えば相手先電話番

号を指示するとコントローラ⑩は、その相手先に回線を繋なぐように電話機を制御する。そして、通話が開始される。なお、この通話の終了時には、音声にて、「ツウワ・シ・ュウ・リ・ヨウ」と発声するか、もしくは、ハンドセットを取り上げる前の状態に戻せば終了となるようになされている。

また、話者が、電話機のハンドセットを取り上げ、専用モードを示す音声、例えば「センヨウ」を、マイクロホン⑪に向かって発声すると、上述した「キョウツウ」と発声した場合と同様にして、パターンマッチング回路⑥で「センヨウ」という単語が認識される。そして、この認識された単語に対応した信号がコントローラ⑩に供給される。すると、コントローラ⑩は、専用モードであることを示す信号を、指紋読取手段⑨および指紋判別手段④に供給する。すると、上述した場合と同様にして、話者の指紋を示す信号が、指紋読取手段⑨から指紋判別手段④に供給される。そして、その指紋を示す信号が登録話者の指紋を示すものであるかどうかが、指紋判別手段④で判別される。

判別された結果、指紋を示す信号が登録話を示すものであれば、その登録話を示す信号をアドレス制御回路⑩を介して専用パターンメモリ(9B)ならびに共通パターンメモリ(9A)に供給する。そして、例えば、第2図に示す共通パターンメモリ(9A)ならびに専用パターンメモリ(9B)の例えば、登録話者Ⅱの単語A、B、Cの音声パターンⅡA、ⅡB、ⅡC、そして、登録話者独自の単語X、Y、Z（例えば、機密である相手先の電話番号を示す暗号）の音声パターンⅡX、ⅡY、ⅡZが選択され、これらが、パターンマッチング回路⑥に供給されるように制御される。そして、上述した場合と同様にして、次にマイクロホン⑪に供給された音声の音声パターンと、標準パターンメモリ⑨から選択された音声パターンとが、パターンマッチング回路⑥で比較される。そして、音声が認識され、その音声に対応した制御がコントローラ⑩にて行なわれる。

もし、指紋読取手段⑨から供給された指紋を示す信号が、登録話者の指紋を示すものではないと、

指紋判別手段④にて判別されると、指紋判別手段④は、それを表す信号をコントローラ⑩に供給する。すると、コントローラ⑩は、音声の受付を停止し、そして、表示装置（図示せず）等に、例えば「センヨウ不可」という表示がなされるようになる。

第4図は、この発明の他の実施例の動作のフローチャートを示す図であり、第1図例と異なる点は、専用モード時の指紋を判別するところである。

図において、ステップ(101)にて、話者がハンドセットを握る。そして、ステップ(102)にて、話者が発した音声を受け付ける。この音声が、共通モードを示すものか、専用モードを示すものかが、ステップ(103)にて、判別される。もし、専用モードでなければ、ステップ(109)にて不特定話者モードと判別され、第1図例に示した、共通パターンメモリ(9A)のみに記憶された音声パターンが選択される。ステップ(103)で専用モードと判別されると、ステップ(104)にて、話者の指紋

が検出される。そして、ステップ(105)にて検出された指紋が登録話者のものであるかが判別され、登録話者のものでなければ、ステップ(110)にて、音声の受け付けが不可とされる。指紋が登録話者のものであれば、ステップ(106)にて、次にその話者が発声した音声を受け付ける。そして、ステップ(107)にて、登録話者の音声認識が、第1図例と同様にして行なわれる。そして、認識された音声に対応した動作がステップ(108)にて行なわれる。

なお、上述した例では、はじめ指紋登録していない話者であっても、電話機を使用できるようになっているが、指紋登録を行なった話者のみが電話機を使用できるようにしてもよい。

なお、指紋登録を行なった話者のみが電話機を使用できるようにした場合には、指紋登録を行なったのみで、音声登録を行なっていない者でも電話機を使用できるようにしてもよい。この場合、指紋登録のみを行なった話者に対しては、不特定話者モードの音声認識が行なわれるものである。

ならびに指紋を登録した登録話者であると判別するようにしたので、登録話者の判別が確実にできる。そして、話者判別をしたときはこの登録話者が登録した音声パターンと入力音声パターンとのパターンマッチングを行なうようにしたので、話者が登録話者である場合の音声認識の応答が速くなる。しかも、登録話者が登録した秘密情報を他人に勝手に使用されないようにすることができるという効果がある。

図面の簡単な説明

第1図はこの発明の一実施例のブロック図、第2図は音声パターンメモリの一例を示す図、第3図は音声登録のフローチャートを示す図、第4図はこの発明の他の実施例の動作のフローチャートを示す図、第5図は音声認識の基本的構成のブロック図である。

(6)はパターンマッチング手段、(7)は指紋読取手段、(8)は指紋判別手段、(9)は標準音声パターンメモリ、(10)は指紋登録手段である。

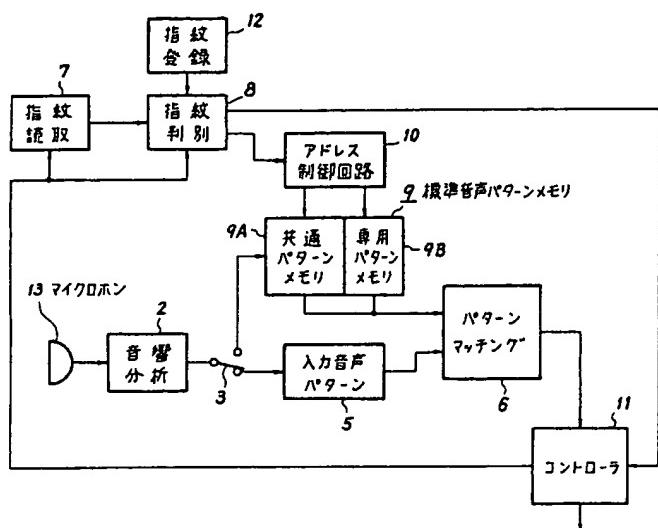
なお、上述した例において、指紋読取手段は、ハンドセットに備えられているが、その他の部分、例えば、ファンクションキーなどが備えられている場合には、そのキーに、指紋読取手段が設けられていてもよいし、また、マイクロホンを用いる場合には、そのマイクロホンの握りの部分に指紋読取手段が設けられていてもよい。

なお、上述した例において、モード判別を行なった後に、指紋の判別が行なわれているが、モード判別を行なう前に、指紋判別を行なうようにしてもよい。

なお上述した実施例は、電話機の場合の例であるが、この発明は電話機だけではなく、例えば、ファクシミリ装置等、音声認識を用いて制御されるものであれば、その他種々のものにも適用することが可能である。

【発明の効果】

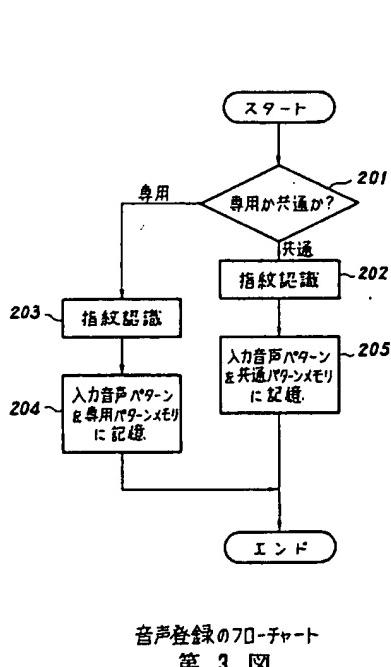
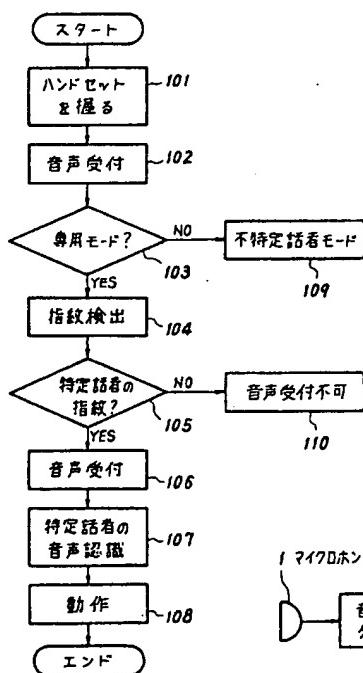
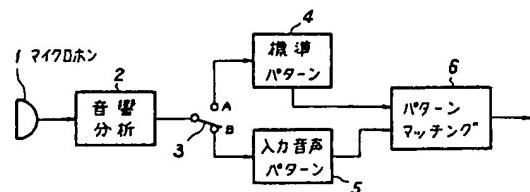
この発明によれば、話者の個人特有のものである指紋の判別を行ない、その話者が音声パターン

実施例
第1図

単語	I	II	III
A	IA	IIA	III A	.	.
B	IB	II B	III B	.	.
C	IC	II C	III C	.	.
⋮	⋮	⋮	⋮	⋮	⋮
X	IX	II X	III X	.	.
Y	IY	II Y	III Y	.	.
Z	I Z	II Z	III Z	.	.

共通パターンメモリ 9A

専用パターンメモリ 9B

音声パターンメモリの一例を示す図
第2図音声登録のフロー・チャート
第3図他の実施例の動作のフロー・チャート
第4図音声認識の基本的構成のアロッカ図
第5図